Locally advanced lung cancer

van Diessen, J.N.A.

2019

document version
Publisher's PDF, also known as Version of record

LOCAL AND REGIONAL TREATMENT RESPONSE BY ^{18}FDG-PET-CT-SCANS 4 WEEKS AFTER CONCURRENT HYPOFRACTIONATED CHEMORADIOTHERAPY IN LOCALLY ADVANCED NSCLC

3

Judi N.A. van Diessen*
Matthew La Fontaine*
Michel M. van den Heuvel
Erik van Werkhoven
Iris Walraven
Wouter V. Vogel
José S.A. Belderbos
Jan-Jakob Sonke

*Both authors contributed equally

Accepted for publication in Radiotherapy and Oncology

Highlights

- A post-treatment FDG-PET-scan at 4 weeks after concurrent chemoradiotherapy in locally advanced NSCLC-patients was beneficial in response evaluation.
- Local and regional treatment response were separately analyzed.
- Associations were found amongst post-treatment and relative response metrics with local/regional/distant failure and overall survival.
- The associations to these outcomes were mostly observed with the primary tumor.
- No preference was found between volumetric and intensity PET-metrics in association to local/regional/distant failure and overall survival.

Background and purpose

To investigate associations of early post-treatment ^{18}Fluorodeoxyglucose-positron-emission-tomography (FDG-PET)-scans with local (LF), regional (RF), distant failure (DF) and overall survival (OS) in locally advanced non-small cell lung cancer (LA-NSCLC)-patients treated with concurrent chemoradiotherapy.

Materials and methods

Forty-seven stage IIIA-B NSCLC-patients included in a randomized phase II-trial (NTR2230) received 66 Gy (24x2.75 Gy) with low dose Cisplatin +/- Cetuximab. FDG-PET-scans were performed at baseline and 4 weeks post-treatment (range, 1.6-10.1). SUV_{max}, SUV_{mean}, metabolic tumor volume (MTV), total lesion glycolysis (TLG) and gross tumor volume were calculated separately for the primary tumor and the involved lymph nodes to generate baseline, post-treatment, and relative response metrics defined as ($metric_{pre}$-$metric_{post}$)/$metric_{pre}$. Univariable cox regression analyses were performed to investigate associations between PET-metrics and outcomes.

Results

Metrics resulted from the post-treatment scan and relative response were associated with outcome, but baseline metrics were not. Primary tumor metrics were stronger associated with all outcomes than lymph node metrics. Both the volumetric (TLG/MTV) and intensity (SUV_{max}/SUV_{mean}) PET-metrics were associated with OS. The intensity metrics were associated with LF, while the volumetric PET-metrics were associated with RF/DF. This was in contrast to the nodal metrics, demonstrating only an association between RF and the relative response of TLG/MTV. No preference was found between PET volumetric and intensity metrics associated with outcome.

Conclusion

Early post-treatment PET-metrics are associated with treatment outcome in LA-NSCLC patients treated with chemoradiotherapy. Both volumetric and intensity PET-metrics are useful, but more for the primary tumor than for lymph nodes.

INTRODUCTION

Improved local control attributes to improved overall survival (OS) in locally advanced non-small cell lung cancer (LA-NSCLC) treated with concurrent chemoradiotherapy (CCRT) [1]. Many adjustments of chemo- and radiotherapy (RT) regimens have been explored to increase outcome [1]. One such adjustment is the combination of daily low dose cisplatin with hypofractionated radiotherapy, which has been explored in several EORTC-trials and in a Dutch randomized phase II Raditux-trial (NTR2230) investigating the addition of Cetuximab to CCRT [2-4]. The results of the Raditux-trial demonstrated a 5-year OS of 37% [5]. Similar results were found in the RTOG-0617 phase 3-trial, claiming that the 5-year OS of 32% is the new benchmark for locally advanced NSCLC [6, 7].

Despite these advancements in CCRT regimens, a substantial share of the patients with locally advanced NSCLC fail either locally or distantly. In order to differentiate patients with a low or high failure risk, investigating biomarkers before and after treatment can be considered. One such prognostic biomarker is imaging of metabolic activity imaging with of ^{18}Fluorodeoxyglucose-positron-emission-tomography(FDG-PET)-scans. FDG-PET has been proven essential in lung cancer for staging, treatment selection and detecting recurrences [8, 9]. However, its role in response measurement is less clear, but may be important for several reasons. First, changes in metabolic activity due to RT are observed sooner than morphologic changes on CT-scans [10, 11]. Second, early metabolic changes characterized by PET-metrics such as the maximum standardized uptake value (SUV_{max}), metabolic tumor volume (MTV) or total lesion glycolysis (TLG) during or short after treatment have been identified as prognostic biomarkers for disease recurrence and survival [8, 12-14]. Thus, an FDG-PET-scan performed shortly after treatment might help to better select patients who might benefit from an additional treatment.

Some questions are still unanswered. Conflicting results have been reported, concerning the timing (during or after) of response FDG-PET-scans [14-16]. Most studies have focused on the primary tumor, while the lymph nodes may merit investigation as well. In the ACRIN/RTOG-0235 study, high residual metabolic activity in lymph nodes 14 weeks after CCRT was associated with a worse locoregional control [17]. Several PET-metrics have been investigated, that can be characterized as volumetric (MTV), intensity metrics (SUV_{max}, SUV_{mean}) or both (TLG) [18, 19]. It is unclear if one of these groups might be preferred in relation to treatment outcome.

The current analysis investigated whether a baseline and a post-treatment FDG-PET-scan at 4 weeks after CCRT was of prognostic value in patients treated in the randomized phase II Raditux-trial. We hypothesized that the early time point of 4 weeks was able to predict OS, local failure (LF), regional failure (RF) and/or distant failure (DF). We investigated also which metric location (primary tumor versus involved lymph nodes) was preferred. Lastly, we tested which PET-metric representation (volumetric versus intensity) could best predict the outcome.

MATERIALS AND METHODS

Study design and patient selection

In the prospective multicenter randomized phase Raditux II trial, 102 patients were treated with CCRT with or without Cetuximab between February 2009 and May 2011 (http://trialregister.nl; Trial ID: NTR2230). The institutional review board approved the trial. The toxicity and long-term OS results were published previously [4, 5].

Diagnostic work-up consisted of a CT-thorax, a whole body FDG-PET-scan<6 weeks before the start of RT, a CT- or Magnetic Resonance Imaging (MRI)-scan of the brain, and pulmonary function tests. Cytology or histology of the PT and/or involved lymph nodes (LNs) to detect the highest N-level was obtained by bronchoscopy, transthoracic biopsy or endoscopic ultrasound of the bronchus (EBUS) or esophagus (EUS).

Patients were scheduled for an early post-treatment FDG-PET-scan in order to select patients, who might be considered for an additional resection. In 72 patients out of 102, an FDG-PET-scan was repeated with a median of 4 weeks (range, 1.6-10.1) after treatment. A CT-thorax was repeated after 6-8 weeks and 24 weeks. Twenty-five patients were excluded: due to adjuvant resection within two months after CCRT (N=18, data in supplementary material), lack of information for scaling the FDG-PET-scan to SUV or gross tumor motion between the PET-emissions and CT-scans hindering registration (N=7).

Chemoradiotherapy

Patients were treated with hypofractionated intensity modulated radiotherapy (IMRT) and concurrent chemotherapy. A four-dimensional planning CT-scan (4DCT) with intravenous contrast was acquired from which a mid-position(MidP)-scan was constructed. The prescribed dose given was 66 Gy in 24 fractions, 5 fractions per week. This was combined with daily low dose Cisplatin (6 mg/m^2, maximum 12 mg), intravenously administered as a bolus 1-2 hours before RT. The total overall treatment time was 32 days.

The gross tumor volume of the primary tumor (GTV_{PT}) was separately delineated from that of the lymph nodes (GTV_{LN}), if possible. Involved LNs were pathologically proven or, in the absence of pathologic evidence, were considered malignant in case of high FDG-uptake and/or growth of LNs on CT. The GTV_{PT} was expanded to a planning target volume (PTV_{PT}) using margins of 12 mm plus 1/4 of the GTV_{PT} peak-to-peak amplitude in orthogonal directions as measured in the 4DCT [20]. An isotropic PTV margin of 12 mm was used for the LNs (PTV_{LN}).

FDG-PET-scans

FDG-PET-scans were made in the treating institution or one of the 4 referring centers. All scans were acquired and reconstructed in accordance with EANM/EARL-guidelines [21]. Patients fasted for 6 hours, and diabetes mellitus needed to be regulated with glucose levels <10 mmol/l. Intravenous administration of FDG was followed by a biodistribution phase of 60+/-10 minutes while resting in a dimly lit room. PET-emission and low dose, free breathing CT without oral or intravenous contrast were acquired from the skull base to the upper thighs. Standard uptake values (SUV) were calculated according to body weight.

Image registration procedure

All delineations for the GTV_{PT} and the GTV_{LN} were performed on the MidP-scan by a dedicated radiation oncologist. For propagation purposes of the delineations to the post-treatment scans, the

post-treatment CT-scan for each patient was rigidly registered to the 4DCT using in-house software. The registration procedure depended on the tumor location: primary tumors without involved LNs were registered according to a tumor match, peripheral tumors with involved LNs according to a bony anatomy match and centrally located tumors with involved LNS according to a carina match. A radiation oncologist visually checked all registrations. The post-treatment CT-scan registrations were then applied to the corresponding post-treatment FDG-PET-emission-scans to align the post-treatment FDG-PET-scans to the 4DCT. The delineations were sampled on both the pre-treatment and post-treatment FDG-PET-scans for imaging analysis.

Image analysis of the FDG-PET-scans

The following PET-metrics were calculated within each region of interest: SUV_{max}, SUV_{mean}, MTV (tumor volume defines as SUV>2.5) and the TLG calculated as SUV_{mean} x MTV. Relative response was calculated by using the formula: $(metric_{pre}\text{-}metric_{post})/metric_{pre}$. These metrics were defined for both the GTV_{PT} and the GTV_{LN}. Although the involved LNs were separately contoured in each patient, we combined them before calculation of the metrics. For SUV_{max}, the maximum intensity was used for the combined lymph nodes. SUV_{mean} was determined by a weighted average of the SUV_{mean} of individual LNs. All other metrics (MTV, TLG) as well as volume (GTV_{LN}) were calculated using the combined sum of the metric for each lymph node (e.g. $MTV_{LN} = MTV_{LN1} + MTV_{LN2} + ...$).

Follow-up and endpoints

The response was evaluated according to Positron Emission Tomography Response Criteria in Solid Tumors (PERCIST 1.0) 4-6 weeks after CCRT. The evaluation of follow-up CT-scans was done according to Response Evaluation Criteria in Solid Tumors (RECIST 1.1). The follow-up schedule was performed according to the study protocol: 6-8 weeks after CCRT followed by 3-monthly visits with physical examination combined with a CT-thorax until disease progression or death. Mortality follow-up was completed up to February 2016.

While the primary endpoint was OS, secondary endpoints were LF, RF and DF. LF was defined as a recurrence or progression of the PT, whereas RF was defined as a recurrence or progression of the treated LNs. In both cases, pathologic confirmation or an increase in tumor diameter of at least 20% compared to the previous CT-scan was scored as a failure (according to RECIST) and sometimes confirmed by PET. DF was defined as a pathological or radiological proof of metastases outside the PTV, a lung tumor in a different ipsilateral lobe, contralateral lobe or any organ outside the lungs.

Statistical analysis

Baseline characteristics are presented as median (+ interquartile range). The follow-up time was calculated from the start of RT and date of death (OS), LF, RF, DF, lost to follow-up or until February 2016. Cumulative incidence-, 6-month-, 1-, 2-, and 5-year mortality rates were investigated. PET-metrics were grouped according to their dissimilarity, which was derived by subtracting the absolute value of the correlation coefficients from one. A dendrogram was created to group metrics by using a dissimilarity threshold of 0.30 (corresponding to a correlation of 0.5 ($r=\pm0.70$). Univariate cox regression analyses were performed to assess associations of each PET-metric grouping with OS, LF, RF and/or DF (noted as the median: [Range] for each group and outcome). Outlier values in the Cox regression were determined and excluded from analysis by using a threshold of 3 times the standard deviation of the score residual [22]. A relative response value of 1 represents a complete response while a value of 0 represents stable disease. The data were analyzed using MATLAB, version R2017b.

RESULTS

Forty-seven patients were selected in the analysis with a median FU of 67 months (IQR: 54-75). Table 1 and 2 depict the patient and tumor characteristics. The median GTV_{PT} and GTV_{LN} were 69.0 cm^3 (IQR: 32.5-147.0) and 28.0 cm^3 (IQR: 12.0-54.0), respectively. In total, 17 patients developed a LF, 9 patients a RF and 24 patients a DF. The 2-year cumulative incidence of LF, RF and DF was 34.3% (N=16), 18.6% (N=9) and 41.9% (N=19), respectively (Fig 1; supplementary material). The median OS was 33 months (95%CI 26-67 months). The 5-year OS rate was 34.8% (95%CI 23.4%-51.9%), consistent with the published 5-year OS rate of the total trial population [5].

Table 1. Patient characteristics of the 47 locally advanced NSCLC-patients treated with concurrent chemoradiotherapy +/-Cetuximab, with a response FDG-PET-scan after a median of 4 weeks.

Patient characteristics		
Median age (year) (IQR)		65 (41-84)
Gender	Male Female	64% 36%
WHO	0 1 2	36% 62% 2%
TNM-stage	IIIA IIIB	60% 40%
T-stage	X 1 2 3 4	2.1% 10.6% 36.2% 14.9% 36.2%
N-stage	0 1 2 3	10.6% 4.3% 59.6% 25.5%
Median volume (cc) (IQR)	GTV_{PT} GTV_{LN}	69.0 (32.5-147.0) 28.0 (12.0-54.0)
Histology	AC SCC LCC	49% 38% 13%
Cetuximab	Yes No	55.3% 44.7%

IQR: interquartile range; GTV_{PT}: gross tumor volume of the primary tumor; GTV_{LN}: gross tumor volume of the involved lymph nodes; SCC: squamous cell carcinoma; AC: adenocarcinoma; LCC: large cell carcinoma;

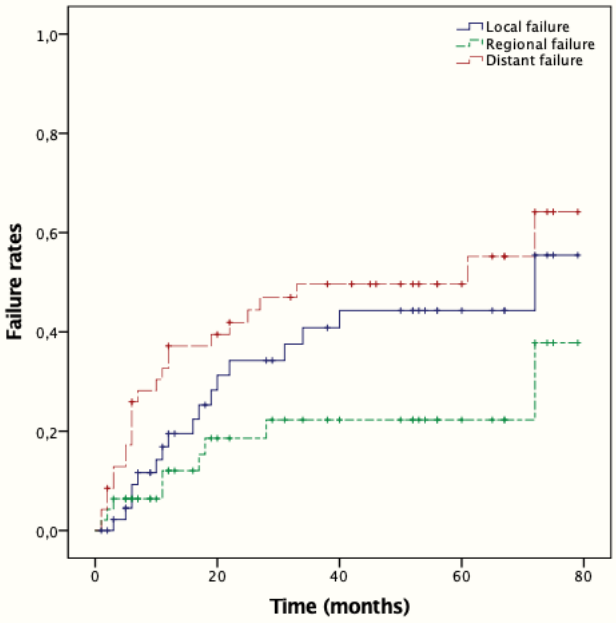
Table 2A-B. Tumor characteristics of the primary tumor (2A) and the lymph nodes (2B) of 47 locally advanced NSCLC-patients treated with concurrent chemoradiotherapy +/-Cetuximab receiving a post-treatment FDG-PET-scan after 4 weeks. The median and interquartile range (IQR) is given of each pre-treatment and post-treatment PET-metric as well as the P-values.

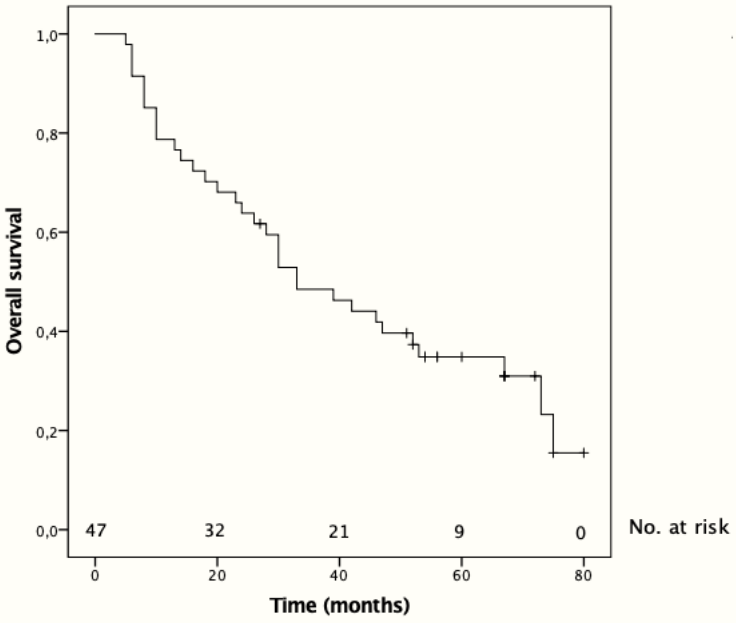
Primary tumor	Pre	Post	P-value	Response (%)
SUV_{max}	13.0 (9.4-16.0)	4.7 (3.5-6.4)	0.002	64.1 (45.3-72.4)
SUV_{mean}	4.1 (3.1-5.8)	1.8 (1.6-2.2)	0.201	58.6 (35.3-69.2)
MTV (cc)	48.3 (21.2-106.7)	10.1 (2.4-35.1)	<0.001	79.0 (47.5-92.7)
TLG (SUV*ml)	292.3 (81.9-697.3)	32.1 (6.5-122.1)	<0.001	89.3 (67.5-97.0)

Lymph nodes	Pre	Post	P-value	Response (%)
SUV_{max}	5.4 (3.7-11.2)	3.8 (2.8-4.5)	0.05	40.0 (8.3-66.6)
SUV_{mean}	2.3 (1.8-3.2)	1.8 (1.6-2.1)	0.014	18.2 (1.6-42.0)
MTV (cc)	7.6 (1.7-26.3)	2.8 (0.2-7.0)	0.042	75.4 (25.5-94.3)
TLG (SUV*ml)	27.2 (5.4-125.5)	7.9 (0.6-19.0)	0.032	78.9 (37.8-96.7)

SUV_{max}: maximum standardized uptake value; SUV_{mean}: mean standardized uptake value; TLG: tumor lesion glycolysis; MTV: metabolic tumor volume; PT: primary tumor; LNs: involved lymph nodes.

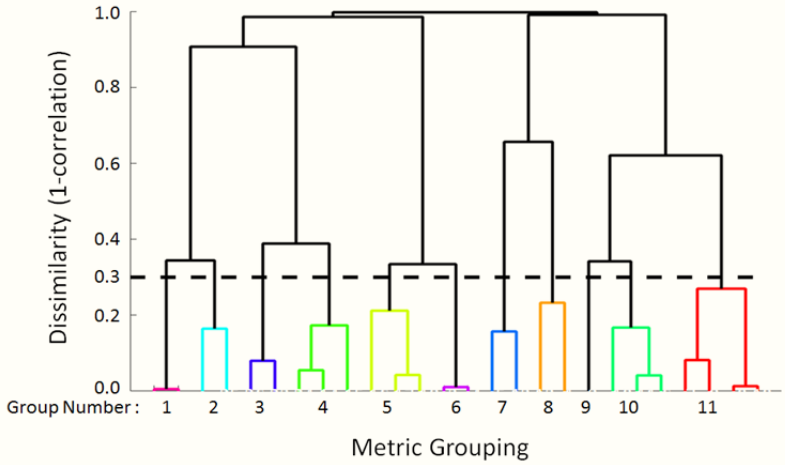
Figure 1. The Kaplan-Meier graphs regarding the LF, RF, DF (1A) and overall survival (1B) of the 47 patients with an early post-treatment FDG-PET-scan, treated with concurrent chemoradiotherapy with or without Cetuximab.





The groupings of the PET-metrics are shown in **Fig. 2**. Although TLG is combination of a volume and intensity metric, our analysis demonstrated a more closely association with MTV than with SUV_{max} and SUV_{mean}. Therefore, TLG and MTV were grouped as volumetric parameters. The hierarchical clustering of the 26 PET-metrics yielded 11 groupings, which were primarily determined by PET-metric location, representation, and type. Of the groupings, five consisted solely of LN-metrics, and 6 solely of primary tumor metrics, with no crossovers within groups. In regards to PET-metric representation, there were four groups of pre-treatment metrics and three for both post-treatment and relative response metrics, plus an additional group mixing GTV post-treatment and response metrics (group 10). PET-metric type was consistently formed of either metric intensity parameters (SUV_{max}, SUV_{mean}) or volumetric parameters (MTV, TLG, GTV_{PT} and GTV_{LN}) with the exception of groups 10 and 11.

Figure 2. PET-metric grouping via hierarchical clustering.



The image above displays the groupings for the PET metrics in a dendrogram (listed in **Table 3**). The PET metrics were grouped by their dissimilarity (y-axis), defined as 1 minus the correlation between metrics in a group. Each group has a maximum dissimilarity of 0.3 (dotted line), translating to a correlation of at least 0.7 amongst all parameters within a group. In total 11 groups were created from the hierarchical clustering.

Table 3. Grouping from left to right in the dendrogram

Group	Site	Type	Metric
1	GTV_{LN}	Response	TLG, MTV
2	GTV_{LN}	Response	SUV_{max}, SUV_{mean}
3	GTV_{LN}	Pre-treatment	SUV_{max}, SUV_{mean}
4	GTV_{LN}	Pre-treatment	TLG, MTV, GTV
5	GTV_{PT}	Pre-treatment	TLG, MTV, GTV
6	GTV_{PT}	Post-treatment	TLG, MTV
7	GTV_{PT}	Pre-treatment	SUV_{max}, SUV_{mean}
8	GTV_{PT}	Response	SUV_{max}, SUV_{mean}
9	GTV_{PT}	Post-treatment	SUV_{max}
10	GTV_{PT}	Post-treatment Response	SUV_{mean} TLG, MTV
11	GTV_{LN}	Post-treatment	SUV_{max}, SUV_{mean}, TLG, MTV

GTV_{PT}: gross tumor volume of the primary tumor; GTV_{LN}: gross tumor volume of the involved lymph nodes; SUV_{max}: maximum standardized uptake value; SUV_{mean}: mean standardized uptake value; TLG: total lesion glycolysis; Response: relative difference between the pre-treatment and post-treatment FDG-PET-scan.

The primary tumor metrics were associated with multiple outcomes. For OS, the post-treatment and the relative response of both the volume and intensity metrics of the primary tumor were highly associated (group 6, 8-10; all P-values <0.05; HR ≥1.01). For LF, the relative response of SUV_{max} and SUV_{mean} of the primary tumor (group 8) was associated (P-value 0.035; HR 1.02). For RF, 4 groups were associated: the post-treatment and relative response of both the intensity and volume metrics of the primary tumor, while the only associated nodal group was the relative response of TLG and MTV (group 1, 6, 9, 10; all P-values <0.05; HR ≥1.003 with a maximum of HR 1.3). For DF, the post-treatment of both TLG and MTV of the primary tumor (group 6) was associated with DF (P-value 0.014; HR 1.008). All associations related to outcome are in Table 4. For the involved lymph nodes, only 1 group was associated with outcome: the relative response of the TLG and MTV (group 1) demonstrated a higher RF rate (HR 1.003). As shown in Table S1C (supplemental), no other group had a median P-value below 0.05 for nodal metrics for any outcome. No parameters from the baseline FDG-PET-scan were associated with any outcome parameter.

Differences in SUV_{max} and SUV_{mean} were compared between patients with an adenocarcinoma (N=23) and squamous cell carcinoma (N=18). We did not find significant differences in the pre-treatment, post-treatment and relative response values (supplemental).

Table 4. The groups showing associations to the outcome according to the median P-value <0.05.

Outcome	Group	Site	Type	Metric	HR (C.I.)	P-value
OS	6	GTV_{PT}	Post	TLG, MTV	1.004 (1.002-1.006) 1.02 (1.007-1.03)	0.00075 0.00061
	9	GTV_{PT}	Post	SUV_{max}	1.17 (1.05-1.31)	0.0053
	8	GTV_{PT}	Response*	SUV_{max}, SUV_{mean}	0.98 (0.97-0.996) 0.98 (0.97-0.999)	0.011 0.040
	10	GTV_{PT}	Post Response*	SUV_{mean} TLG, MTV	1.65 (0.98-2.76) 0.987 (0.976-0.998) 0.988 (0.978-0.999)	0.058 0.021 0.027
LF	8	GTV_{PT}	Response*	SUV_{max}, SUV_{mean}	0.98 (0.96-0.997) 0.98 (0.96-0.9995)	0.024 0.045
RF	9	GTV_{PT}	Post	SUV_{max}	1.32 (1.07-1.63)	0.0097
	6	GTV_{PT}	Post	TLG, MTV	1.004 (1.0009-1.008) 1.02 (1.004-1.03)	0.014 0.014
	1	GTV_{LN}	Response*	TLG, MTV	0.997 (0.995-0.9998) 0.996 (0.993-0.999)	0.056 0.012
	10	GTV_{PT}	Post Response*	SUV_{mean} TLG, MTV	2.45 (1.04-5.8) 0.99 (0.97-1.02) 0.98 (0.96-1.00004)	0.04 0.61 0.05
DF	6	GTV_{PT}	Post	TLG, MTV	1.003 (1.0006-1.006) 1.01 (1.003-1.02)	0.014 0.013

OS: overall survival; LF: local failure; RF: regional failure; DF: distant failure.
*A greater response indicates a beneficial or protective effect.

DISCUSSION

This study investigated the prognostic value of a baseline and an early post-treatment FDG-PET-scan at 4 weeks with outcomes of stage IIIA-B NSCLC-patients treated with CCRT in the phase II Raditux-trial. The results demonstrated that a post-treatment FDG-PET-scan was associated with LF, RF, DF and OS. In addition, we investigated association to outcome based on assessment location (primary tumor versus involved lymph nodes), which showed most associations were observed for the primary tumor. Lastly, we did not find a preference between PET volumetric and intensity metrics in association to outcome, demonstrating a lack of a superior group of metrics. In contrast, the GTV_{PT} and/or GTV_{LN} were not associated with outcome, but the metabolic volume (MTV and TLG) were associated with OS, RF and DF, demonstrating the potential benefit of post-treatment FDG-PET-scans in the response evaluation.

Our first research question concerned the usefulness of a post-treatment FDG-PET-scan. We demonstrated that the post-treatment and the relative response of both the volumetric and intensity PET-metrics were associated with OS, LF, RF and DF. This finding is consistent with literature. Jeong et al. reported the results of 119 stage III NSCLC-patients treated with CCRT, who received an FDG-PET-scan after 2-4 weeks [23]. The multivariable analysis showed a significant association of the relative response SUV_{max} of the primary tumor with OS. Additionally, Usmanji et al. found in 28 stage IIIA-B NSCLC-patients, treated with CCRT and receiving a FDG-PET-scan after 2 weeks and 3 months, a significant association between the relative response TLG and progression-free survival [12]. In the current analysis, we did not find an association between the pre-treatment FDG-PET-scan and LF/RF, although this was demonstrated in literature [19, 24]. One possible explanation why our study did not find similar results is due to the limited number of patients analyzed and the exclusion of 18 patients that were selected for additional surgery. We tested whether the exclusion of the resected patients could have affected the results by comparing prognostic factors. We did not find differences that might explain why there were no associations with the pre-treatment FDG-PET-scan (supplementary material). Another explanation might be the interval of maximum six weeks between the pre-treatment FDG-PET-scan and the radiotherapy, which is rather long. Everitt et al. have described a progression probability of 29% in three weeks without treatment [25]. The recently published ESTRO ACROP guidelines advise a maximum interval of three weeks between the FDG-PET-scan and the start of treatment [26].

Our study demonstrated that performing an FDG-PET-scan at an early time point of 4 weeks after CCRT has prognostic value. This is in line with the previously discussed results by Jeong et al. and Usmanji et al [12, 23]. In addition, the non-randomized phase II ACRIN-6668/RTOG-0235 demonstrated similar findings, although the time point of the post-treatment FDG-PET-scans was at a median of 14 weeks [16]. This trial investigated 173 stage III NSCLC-patients with inoperable NSCLC treated with CCRT. A high post-treatment SUV_{max} of both GTV_{PT} and GTV_{LN} was significantly associated with a poor OS, although a clear cut-off value was not found. Our analysis demonstrated similar results at an earlier time point of 4 weeks after CCRT, which allows clinicians to decide for intensified additional treatments. Earlier time points e.g. during treatment might also be of interest, although the purpose might be primarily to adapt the radiation therapy. This was recently investigated by Gensheimer et al. and Kong et al. [14, 27].

Our second hypothesis was that the metric location might play a decisive role in relation to outcome. Although the patient population is small, the number of patients with involved lymph nodes (N=42) does not differ greatly to the number of patients with a primary tumor (N=46). We

demonstrated, by separately analyzing the GTV_{PT} and the $GTV_{LN,}$ that the primary tumor was the location with the strongest associations. Most studies analyzed the GTV as a combination of GTV_{PT} and GTV_{LN} [14, 23, 27]. Bissonnette et al. separately analyzed the primary tumor and involved lymph nodes as well, but performed FDG-PET-scans during CCRT at 0, 2, 4 and 7 weeks in 27 locally advanced NSCLC-patients [28]. The authors observed significant associations of nodal PET-metrics with OS, which contrasted with our results. Our analysis demonstrated only associations of volumetric nodal metrics with RF. We did not find additional associations of nodal PET-metrics. This might be explained by the fact that all LNs from 1 patient were jointly analyzed for statistical purposes, although each LN was contoured separately. The LNs with the highest SUV_{max} on the pre-treatment and post-treatment FDG-PET-scan were chosen, which were not necessarily the same. Future studies may be able to include more patients, which will provide statistical power to analyze the LN stations separately as the predictive power of all LNs together may not be the same.

Lastly, we were interested in differences between the volumetric (MTV and TLG) and the intensity (SUV_{max} and SUV_{mean}) metrics. We found no preference since both groups were associated with outcome. This is also reflected in literature; Jeong et al. demonstrated the relevance of SUV_{max} whereas others favored the metabolic volume [12, 23]. However, we did not demonstrate an association of the GTV_{PT} and/or GTV_{LN}. Therefore we support the use of post-treatment FDG-PET-scans at an early time point by demonstrating in the current analysis that the metabolic response is able to inform us about the prognosis. We haven't been able to find a practical PET-metric that might be used in future analyses. Interpretation as well as extrapolation of the results may also be hampered by the hypofractionated schedule resulting in a reduced overall treatment time and a higher radiation dose. Nevertheless, based on our results we may suggest focusing on the primary tumor only since associations with involved lymph nodes were minimal.

Hierarchical clustering was able to differentiate independent tumor metrics and provide consistent outcome analysis. The hierarchical clustering of the dendrogram demonstrated the complementary information of metric location (GTV_{PT} vs. GTV_{LN}), metric representation (pre-treatment, post-treatment, response), and metric type (intensity, volume), as only two of eleven groups had some intermixing of time point or metric type. In contrast, metrics of the same type (e.g. intensity: SUV_{max}, SUV_{mean}) were found to contribute similar information regardless of metric location or metric time point as they were reliably grouped together. Furthermore, the associations with outcome of parameters within a cluster were consistent p-values, demonstrating a median range within a group of 0.09 [IQR: 0.02-0.25]. Groups with weaker associations may have been impacted by the low number of events in RF (n = 9), differences in the number of outliers within a group, or ill-defined and value capped responses (Group 11).

There are limitations to this analysis. First, the timeframe of the post-treatment FDG-PET-scans was rather wide (range, 1.6-10.1) causing the results less comparable with each other. Second, the number of included patients is limited, as previously mentioned. Due to the small number of events, the study was susceptible to outliers impacting the associations to outcome. Investigation of the outliers for the primary tumor showed that up to 4 patients affected the associations to outcome for each metric type, displaying either a high pre- or post-treatment value, or minimal response. While these characteristics are often associated with poor outcome in literature, these patients had a favorable outcome [10, 16]. As these patients had high parameter values in comparison to the rest of the study, they had a strong influence in the Cox regression residuals. In comparison, there were no outliers for the lymph nodes, which had strong associations to outcome. Third, all patients were included in a randomized phase II trial with 98% having a good performance status (WHO 0-1), which

partly explains the high OS. Fourth, we excluded 18 patients that received an adjuvant resection, which might have affected the failure and survival rates as well as its associations. Fifth, the analysis in this study consisted of multiple statistical tests, which increases the probability of type I errors. Since many PET-metrics were tested (although the groupings reduced this considerably), one could apply a correction for multi-testing. While we focused on metrics with $p<0.05$, there was no correction for multiple testing or even for significance testing, leading to this paper being hypothesis-generating in nature. The last limitation of this analysis is that more than half of the pre-treatment FDG-PET-scans were performed in different hospitals, while all post-treatment FDG-PET-scans were performed in 1 hospital. This might explain why we did not find associations with the pre-treatment FDG-PET-scan. Although the EANM/EARL-guidelines were followed, the calculation of the PET-metrics is influenced by the machine equipment and local hospital protocols and might have impacted the GTV_{PT} and GTV_{LN}.

In conclusion, in patients with locally advanced stage NSCLC treated with concurrent chemoradiotherapy, a post-treatment FDG-PET-scan at an early time point of 4 weeks demonstrated prognostic value, depending on the applied metric and evaluated tumor location. The primary tumor showed associations of both volumetric and intensity PET-metrics with LF, RF, DF and OS, while the involved lymph nodes only showed associations with RF. Both the volumetric and intensity PET-metrics were equally useful.

Funding

This research was partially funded by Health Holland public-private partnership grant (LSHM15036) in collaboration with Elekta Oncology Systems AB.

REFERENCES

[1] Auperin A, Le Pechoux C, Rolland E, Curran WJ, Furuse K, Fournel P, et al. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer. J Clin Oncol. 2010;28:2181-90.

[2] Schaake-Koning C, van den Bogaert W, Dalesio O, Festen J, Hoogenhout J, van Houtte P, et al. Effects of concomitant cisplatin and radiotherapy on inoperable non-small-cell lung cancer. N Engl J Med. 1992;326:524-30.

[3] Belderbos J, Uitterhoeve L, van Zandwijk N, Belderbos H, Rodrigus P, van de Vaart P, et al. Randomised trial of sequential versus concurrent chemo-radiotherapy in patients with inoperable non-small cell lung cancer (EORTC 08972-22973). Eur J Cancer. 2007;43:114-21.

[4] van den Heuvel MM, Uyterlinde W, Vincent AD, de Jong J, Aerts J, Koppe F, et al. Additional weekly Cetuximab to concurrent chemoradiotherapy in locally advanced non-small cell lung carcinoma: efficacy and safety outcomes of a randomized, multi-center phase II study investigating. Radiother Oncol. 2014;110:126-31.

[5] Walraven I, van den Heuvel M, van Diessen J, Schaake E, Uyterlinde W, Aerts J, et al. Long-term follow-up of patients with locally advanced non-small cell lung cancer receiving concurrent hypofractionated chemoradiotherapy with or without cetuximab. Radiother Oncol. 2016;118:442-6.

[6] Bradley JD, Paulus R, Komaki R, Masters G, Blumenschein G, Schild S, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): a randomised, two-by-two factorial phase 3 study. Lancet Oncol. 2015;16:187-99.

[7] Bradley JD, Hu C, Komaki RU, Masters G, Blumenschein GR, Schild SE, et al. Long-Term Results of RTOG 0617: A Randomized Phase 3 Comparison of Standard Dose Versus High Dose Conformal Chemoradiation Therapy +/- Cetuximab for Stage III NSCLC. International Journal of Radiation Oncology • Biology • Physics. 2017;99:S105.

[8] Grootjans W, de Geus-Oei LF, Troost EG, Visser EP, Oyen WJ, Bussink J. PET in the management of locally advanced and metastatic NSCLC. Nat Rev Clin Oncol. 2015;12:395-407.

[9] He YQ, Gong HL, Deng YF, Li WM. Diagnostic efficacy of PET and PET/CT for recurrent lung cancer: a meta-analysis. Acta radiologica. 2014;55:309-17.

[10] Grootjans W, Usmanij EA, Oyen WJ, van der Heijden EH, Visser EP, Visvikis D, et al. Performance of automatic image segmentation algorithms for calculating total lesion glycolysis for early response monitoring in non-small cell lung cancer patients during concomitant chemoradiotherapy. Radiother Oncol. 2016;119:473-9.

[11] Cheebsumon P, Boellaard R, de Ruysscher D, van Elmpt W, van Baardwijk A, Yaqub M, et al. Assessment of tumour size in PET/CT lung cancer studies: PET- and CT-based methods compared to pathology. EJNMMI Res. 2012;2:56.

[12] Usmanij EA, de Geus-Oei LF, Troost EG, Peters-Bax L, van der Heijden EH, Kaanders JH, et al. 18F-FDG PET early response evaluation of locally advanced non-small cell lung cancer treated with concomitant chemoradiotherapy. J Nucl Med. 2013;54:1528-34.

[13] Hyun SH, Ahn HK, Kim H, Ahn MJ, Park K, Ahn YC, et al. Volume-based assessment by (18)F-FDG PET/CT predicts survival in patients with stage III non-small-cell lung cancer. Eur J Nucl Med Mol Imaging. 2014;41:50-8.

[14] Gensheimer MF, Hong JC, Chang-Halpenny C, Zhu H, Eclov NCW, To J, et al. Mid-radiotherapy PET/CT for prognostication and detection of early progression in patients with stage III non-small cell lung cancer. Radiother Oncol. 2017.

[15] Cremonesi M, Gilardi L, Ferrari ME, Piperno G, Travaini LL, Timmerman R, et al. Role of interim (18)F-FDG-PET/CT for the early prediction of clinical outcomes of Non-Small Cell Lung Cancer (NSCLC) during radiotherapy or chemo-radiotherapy. A systematic review. Eur J Nucl Med Mol Imaging. 2017;44:1915-27.

[16] Machtay M, Duan F, Siegel BA, Snyder BS, Gorelick JJ, Reddin JS, et al. Prediction of survival by [18F] fluorodeoxyglucose positron emission tomography in patients with locally advanced non-small-cell lung cancer undergoing definitive chemoradiation therapy: results of the ACRIN 6668/RTOG 0235 trial. J Clin Oncol. 2013;31:3823-30.

[17] Markovina S, Duan F, Snyder BS, Siegel BA, Machtay M, Bradley JD. Regional Lymph Node Uptake of [(18)F]Fluorodeoxyglucose After Definitive Chemoradiation Therapy Predicts Local-Regional Failure of Locally Advanced Non-Small Cell Lung Cancer: Results of ACRIN 6668/RTOG 0235. Int J Radiat Oncol Biol Phys. 2015;93:597-605.

[18] Salavati A, Duan F, Snyder BS, Wei B, Houshmand S, Khiewvan B, et al. Optimal FDG PET/CT volumetric parameters for risk stratification in patients with locally advanced non-small cell lung cancer: results from the ACRIN 6668/RTOG 0235 trial. Eur J Nucl Med Mol Imaging. 2017;44:1969-83.

[19] Paesmans M, Garcia C, Wong CY, Patz EF, Jr., Komaki R, Eschmann S, et al. Primary tumour standardised uptake value is prognostic in nonsmall cell lung cancer: a multivariate pooled analysis of individual data. Eur Respir J. 2015;46:1751-61.

[20] van Herk M, Remeijer P, Lebesque JV. Inclusion of geometric uncertainties in treatment plan evaluation. Int J Radiat Oncol Biol Phys. 2002;52:1407-22.

[21] Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging. 2015;42:328-54.

[22] Therneau TM, Grambsch, P. and Fleming T.R. Martingale-based residuals for survival models. https://wwwjstororg/stable/pdf/2336057pdf 1990.

[23] Jeong JU, Chung WK, Nam TK, Song JY, Yoon MS, Kim YC, et al. Early metabolic response on 18F-fluorodeoxyglucose-positron-emission tomography/computed tomography after concurrent chemoradiotherapy for advanced stage III non-small cell lung cancer is correlated with local tumor control and survival. Anticancer research. 2014;34:2517-23.
[24] Na F, Wang J, Li C, Deng L, Xue J, Lu Y. Primary tumor standardized uptake value measured on F18-Fluorodeoxyglucose positron emission tomography is of prediction value for survival and local control in non-small-cell lung cancer receiving radiotherapy: meta-analysis. J Thorac Oncol. 2014;9:834-42.
[25] Kong FS, Li L, Wang W, Campbell J, Waller JL, Piert M, et al. Greater reduction in mid-treatment FDG-PET volume may be associated with worse survival in non-small cell lung cancer. Radiother Oncol. 2019;132:241-9.
[26] Bissonnette JP, Yap ML, Clarke K, Shessel A, Higgins J, Vines D, et al. Serial 4DCT/4DPET imaging to predict and monitor response for locally-advanced non-small cell lung cancer chemo-radiotherapy. Radiother Oncol. 2018;126:347-54.

SUPPLEMENTARY MATERIAL

Figure S1. The incidence of LF, RF and DF of the 47 locally advanced NSCLC-patients treated with concurrent chemoradiotherapy +/- Cetuximab (crude percentages given, i.e. events at data cutoff as percentage of total number of patients).

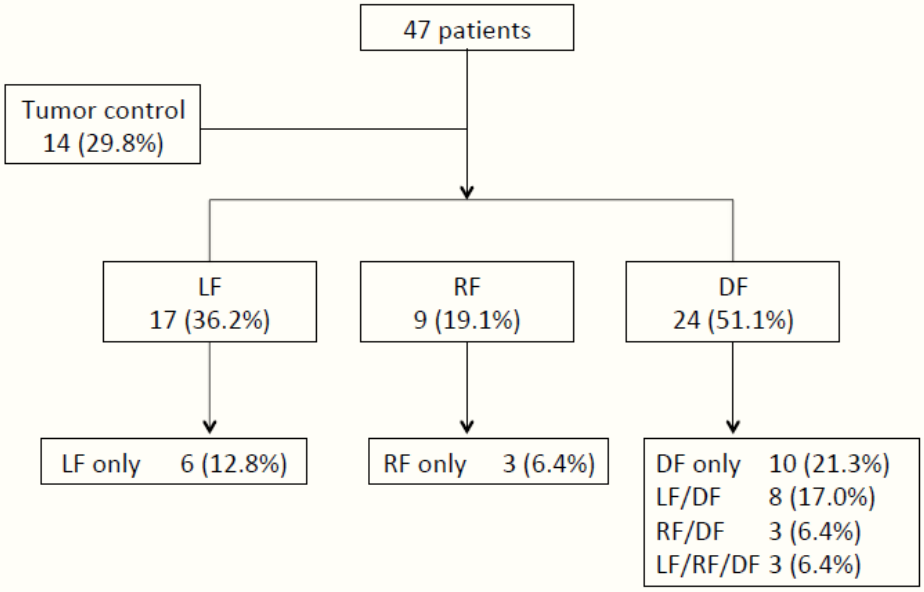


Table S1A. Associations of the grouped PET-metrics to overall survival. Groups in ascending order according to the median P-value.

Group	Site	Type	Metric	HR (range)	P-value (range)
6	GTV_{PT}	Post-treatment	TLG, MTV	1.01 (1.004-1.016)	0.001 (0.001-0.001)
9	GTV_{PT}	Post-treatment	SUV_{max}	1.174	0.005
8	GTV_{PT}	Response	SUV_{max}, SUV_{mean}	1.017 (1.017-1.018)	0.026 (0.011-0.04)
10	GTV_{PT}	Post-treatment Response	SUV_{mean} TLG, MTV	1.013 (1.012-1.648)	0.027 (0.022-0.058)
11	GTV_{LN}	Post-treatment	SUV_{max}, SUV_{mean}, TLG, MTV	1.029 (1.006-1.518)	0.111 (0.065-0.733)
5	GTV_{PT}	Pre-treatment	TLG, MTV, GTV	1.003 (1.001-1.005)	0.159 (0.149-0.159)
7	GTV_{PT}	Pre-treatment	SUV_{max}, SUV_{mean}	1.09 (1.041-1.139)	0.166 (0.157-0.176)
2	GTV_{LN}	Response	SUV_{max}, SUV_{mean}	1.006 (1.004-1.007)	0.227 (0.165-0.289)
1	GTV_{LN}	Response	TLG, MTV	1.001 (1.001-1.002)	0.38 (0.378-0.382)
3	GTV_{LN}	Pre-treatment	SUV_{max}, SUV_{mean}	0.945 (0.933-0.957)	0.455 (0.106-0.805)
4	GTV_{LN}	Pre-treatment	TLG, MTV, GTV	1 (0.999-1.003)	0.898 (0.708-0.933)

Table S1B. Associations of the grouped PET-metrics to local failure. Groups in ascending order according to the median P-value.

Group	Site	Type	Metric	HR (range)	P-value (range)
8	GTV_{PT}	Response	SUV_{max}, SUV_{mean}	1.023 (1.023-1.023)	0.035 (0.024-0.045)
10	GTV_{PT}	Post-treatment Response	SUV_{mean} TLG, MTV	1.014 (1.014-1.656)	0.096 (0.062-0.154)
7	GTV_{PT}	Pre-treatment	SUV_{max}, SUV_{mean}	0.828 (0.764-0.892)	0.11 (0.07-0.149)
3	GTV_{LN}	Pre-treatment	SUV_{max}, SUV_{mean}	0.747 (0.597-0.897)	0.115 (0.11-0.119)
1	GTV_{LN}	Response	TLG, MTV	1.002 (1.002-1.002)	0.187 (0.175-0.198)
4	GTV_{LN}	Pre-treatment	TLG, MTV, GTV	0.989 (0.95-0.996)	0.247 (0.154-0.339)
6	GTV_{PT}	Post-treatment	TLG, MTV	1.005 (1.002-1.009)	0.286 (0.225-0.347)
2	GTV_{LN}	Response	SUV_{max}, SUV_{mean}	1.006 (1.006-1.006)	0.287 (0.167-0.406)
5	GTV_{PT}	Pre-treatment	TLG, MTV, GTV	0.998 (0.995-0.999)	0.329 (0.239-0.484)
11	GTV_{LN}	Post-treatment	SUV_{max}, SUV_{mean}, TLG, MTV	0.863 (0.631-0.998)	0.479 (0.301-0.509)
9	GTV_{PT}	Post-treatment	SUV_{max}	1.023	0.82

Table S1C. Associations of the grouped PET-metrics to regional failure. Groups in ascending order according to the median P-value.

Group	Site	Type	Metric	HR (range)	P-value (range)
9	GTV_{PT}	Post-treatment	SUV_{max}	1.322	0.001
6	GTV_{PT}	Post-treatment	TLG, MTV	1.011 (1.004-1.018)	0.014 (0.014-0.014)
1	GTV_{LN}	Response	TLG, MTV	1.003 (1.003-1.004)	0.034 (0.012-0.056)
10	GTV_{PT}	Post-treatment Response	SUV_{mean} TLG, MTV	1.022 (1.007-2.453)	0.05 (0.04-0.608)
2	GTV_{LN}	Response	SUV_{max}, SUV_{mean}	1.015 (1.015-1.016)	0.067 (0.021-0.112)
5	GTV_{PT}	Pre-treatment	TLG, MTV, GTV	1.005 (1.002-1.011)	0.075 (0.025-0.146)
4	GTV_{LN}	Pre-treatment	TLG, MTV, GTV	0.974 (0.954-0.989)	0.249 (0.171-0.253)
7	GTV_{PT}	Pre-treatment	SUV_{max}, SUV_{mean}	1.171 (1.163-1.179)	0.262 (0.001-0.523)
3	GTV_{LN}	Pre-treatment	SUV_{max}, SUV_{mean}	0.813 (0.765-0.861)	0.312 (0.178-0.446)
11	GTV_{LN}	Post-treatment	SUV_{max}, SUV_{mean}, TLG, MTV	1.298 (1.005-3.719)	0.321 (0.002-0.705)
8	GTV_{PT}	Response	SUV_{max}, SUV_{mean}	1.013 (1.007-1.019)	0.466 (0.263-0.668)

Table S1D. Associations of the grouped PET-metrics to distant failure. Groups in ascending order according to the median P-value.

Group	Site	Type	Metric	HR (range)	P-value (range)
6	GTV_{PT}	Post-treatment	TLG, MTV	1.008 (1.003-1.013)	0.014 (0.013-0.014)
2	GTV_{LN}	Response	SUV_{max}, SUV_{mean}	1.01 (1.008-1.011)	0.068 (0.033-0.104)
5	GTV_{PT}	Pre-treatment	TLG, MTV, GTV	1.004 (1-1.007)	0.087 (0.054-0.688)
10	GTV_{PT}	Post-treatment Response	SUV_{mean} TLG, MTV	1.012 (1.01-1.238)	0.101 (0.081-0.299)
8	GTV_{PT}	Response	SUV_{max}, SUV_{mean}	1.013 (1.013-1.014)	0.109 (0.069-0.15)
7	GTV_{PT}	Pre-treatment	SUV_{max}, SUV_{mean}	0.898 (0.851-0.946)	0.203 (0.165-0.242)
3	GTV_{LN}	Pre-treatment	SUV_{max}, SUV_{mean}	0.879 (0.847-0.911)	0.222 (0.079-0.365)
9	GTV_{PT}	Post-treatment	SUV_{max}	1.065	0.364
1	GTV_{LN}	Response	TLG, MTV	1.001 (1.001-1.001)	0.445 (0.427-0.462)
11	GTV_{LN}	Post-treatment	SUV_{max}, SUV_{mean}, TLG, MTV	0.989 (0.902-1.011)	0.448 (0.302-0.944)
4	GTV_{LN}	Pre-treatment	TLG, MTV	1.005 (1.001-1.008)	0.503 (0.402-0.658)

Table S2. The cumulative incidence of LF, RF and DF of the 47 locally advanced NSCLC-patients treated with concurrent chemoradiotherapy +/- Cetuximab. Additionally, the OS rates are demonstrated.

	1-year	2-year	3-year	5-year
LF	19.5%	34.3%	40.8%	44.3%
RF	12.1%	18.6%	22.3%	22.3%
DF	32.7%	41.9%	49.6%	49.6%
OS	78.7%	63.8%	48.5%	34.8%

LF: local failure; RF: regional failure; DF: distant failure; OS: overall survival

Table S3. Comparison of the SUV_{max}, SUV_{mean}, TLG and MTV between the patients with an adenocarcinoma and squamous cell carcinoma. No significant differences were observed.

	SUVmax			SUVmean		
	Pre-treatment	Post-treatment	Response	Pre-treatment	Post-treatment	Response
Primary tumor	0.340	0.751	0.888	0.751	0.751	0.340
Lymph nodes	0.848	0.135	0.848	0.940	0.135	0.848

	TLG			MTV		
	Pre-treatment	Post-treatment	Response	Pre-treatment	Post-treatment	Response
Primary tumor	0.112	0.751	0.340	0.112	0.340	0.751
Lymph nodes	0.848	0.135	0.507	0.848	0.135	0.646

Table S4. Comparison of the 47 patients treated with chemoradiotherapy (cCRT) versus the 18 patients treated with chemoradiotherapy followed by resection (Trimodality).

Characteristic (Median;IQR)	cCRT	Trimodality	P-value
GTV_{PT}	69.0 (32.5-147.0)	68.0 (45.8-100.0)	0.676
GTV_{LN}	28.0 (12.0-54.0)	17.0(10.5-111.8)	0.561
Pre-treatment SUV_{max} PT	12.9 (9.3-15.7)	14.1 (10.1-15.9)	0.738
Post-treatment SUV_{max} PT	4.8 (3.5-6.4)	9.3 (3.2-12.6)	0.215
Pre-treatment SUV_{max} LN	5.4 (3.7-11.1)	5.8 (3.0-9.6)	0.318
Post-treatment SUV_{max} LN	3.8 (2.8-4.5)	2.8 (2.4-4.9)	0.076
OS (Median)	33.0	56.0	0.385

Comparison of the resected and non-resected patients

The median GTV_{PT} and the GTV_{LN} of the resected and non-resected group did not differ significantly. In addition, the median pre- and post-treatment SUV_{max} of the primary tumor and the lymph nodes were comparable between the 2 groups and both patient groups had a good performance status. However, the median OS seemed more favorable in the patients with an adjuvant resection: 56 months versus 33 months for the patients treated with CCRT, although it was not significantly different (p=0.385). The pathologic report after the resection revealed residual microscopic or macroscopic tumor in 12 of the 18 patients, which may explain the better OS.

Figure S2. The overall survival of the 47 patients treated with chemoradiotherapy (cCRT) versus the 18 patients treated with chemoradiotherapy followed by resection (Trimodality).